

Supplementary Material - Computational Approaches in Chemical Space Exploration for Carbon Fixation Pathways

Anne-Susann Abel, Nino Lauber, Jakob Lykke Andersen, Rolf Fagerberg, Daniel Merkle, and Christoph Flamm

Table S1: Collection of the 49 initial molecules for the chemical space expansion, with their molecule name as well as their input format. Some molecules have abstract labels like [CoA] instead of atom labels for ease of use. CoA: coenzyme A, 3HP: 3-hydroxypropionic acid, 4HB: 4-hydroxybutyric acid, PEP: phosphoenolpyruvic acid, GAP: glyceraldehyde-3-phosphate, DHAP: dihydroxyacetone phosphate.

Molecule Name	Smiles Representation
SUCCINYL COA	<chem>OC(=O)CCC(=O)S[CoA]</chem>
SUCCINATE SEMIALDEHYDE	<chem>O=CCCC(=O)O</chem>
4HB	<chem>C(CC(=O)O)CO</chem>
4-HYDROXYBUTYRYL COA	<chem>C(CC(=O)S[CoA])CO</chem>
CROTONYL-COA	<chem>CC=CC(=O)S[CoA]</chem>
3-HYDROXYBUTYRYL COA	<chem>CC(CC(=O)S[CoA])O</chem>
ACETOACETYL COA	<chem>CC(=O)CC(=O)S[CoA]</chem>
ACETYL COA	<chem>CC(=O)S[CoA]</chem>
PYRUVATE	<chem>CC(=O)C(=O)O</chem>
PEP	<chem>C=C(C(=O)O)OP(=O)(O)O</chem>
OXALACETIC ACID	<chem>C(C(=O)C(=O)O)C(=O)O</chem>
MALIC ACID	<chem>C(C(C(=O)O)O)C(=O)O</chem>
FUMARIC ACID	<chem>C(=CC(=O)O)C(=O)O</chem>

SUCCINIC ACID	<chem>OC(=O)CCC(=O)O</chem>
MALONYL COA	<chem>C(C(=O)O)C(=O)S[CoA]</chem>
MALONATE SEMIALDEHYDE	<chem>OC(=O)CC=O</chem>
3HP	<chem>C(CO)C(=O)O</chem>
3-HYDROXYPROPIONATE COA	<chem>C(CO)C(=O)S[CoA]</chem>
ACRYLYL COA	<chem>C=CC(=O)S[CoA]</chem>
METHYLMALONYL COA	<chem>CC(C(=O)O)C(=O)S[CoA]</chem>
2-KETOGLUTARIC ACID	<chem>C(CC(=O)O)C(=O)C(=O)O</chem>
ISOCITRIC ACID	<chem>C(C(C(C(=O)O)O)C(=O)O)C(=O)O</chem>
ACONITIC ACID	<chem>C(C(=CC(=O)O)C(=O)O)C(=O)O</chem>
CITRIC ACID	<chem>C(C(=O)O)C(CC(=O)O)(C(=O)O)O</chem>
CITRYL COA	<chem>C(C(=O)O)C(CC(=O)S[CoA])(C(=O)O)O</chem>
MALYL COA	<chem>C(C(C(=O)O)O)C(=O)S[CoA]</chem>
GLYOXYLATE	<chem>C(=O)C(=O)O</chem>
PROPIONYL COA	<chem>CCC(=O)S[CoA]</chem>
METHYLMALYL COA	<chem>CC(C(C(=O)O)O)C(=O)S[CoA]</chem>
MESACONYL C1 COA	<chem>CC(=CC(=O)O)C(=O)S[CoA]</chem>
MESACONYL C4 COA	<chem>CC(=CC(=O)S[CoA])C(=O)O</chem>
CITRAMALYL COA	<chem>CC(CC(=O)S[CoA])(C(=O)O)O</chem>
ETHYLMALONYL COA	<chem>CCC(C(=O)O)C(=O)S[CoA]</chem>
METHYLSUCCINYL COA	<chem>OC(=O)CC(C)C(=O)S[CoA]</chem>
RIBULOSE-1_5-BISPHOSPHATE	<chem>C(C(C(C(=O)COP(=O)(O)O)O)OP(=O)(O)O</chem>
3-PHOSPHOGLYCERATE	<chem>C(C(C(=O)O)O)OP(=O)(O)O</chem>
1_3-BISPHOSPHOGLYCERATE	<chem>C(C(C(=O)OP(=O)(O)O)OP(=O)(O)O</chem>
GAP	<chem>C(C(C(=O)O)OP(=O)(O)O</chem>
DHAP	<chem>C(C(=O)COP(=O)(O)O)O</chem>

FRUCTOSE-1_6-BISPHOSPHATE	<chem>C(C(C(C(C(=O)COP(=O)(O)O)O)O)OP(=O)(O)O</chem>
FRUCTOSE-6-P	<chem>C(C(C(C(C(=O)CO)O)O)OP(=O)(O)O</chem>
ERYTHROSE-4-P	<chem>C(C(C(C(=O)O)O)OP(=O)(O)O</chem>
RIBULOSE-5-P	<chem>C(C(C(C(C(=O)CO)O)O)OP(=O)(O)O</chem>
XYLULOSE-5-P	<chem>C(C(C(C(C(=O)CO)[xyO][H])O)OP(=O)(O)O</chem>
SEDOHEPTULOSE-1_7-BISPHOSPHATE	<chem>C(C(C(C(C(C(=O)COP(=O)(O)O)O)O)OP(=O)(O)O</chem>
SEDOHEPTULOSE-7-PHOSPHATE	<chem>C(C(C(C(C(C(=O)CO)O)O)OP(=O)(O)O</chem>
RIBOSE-5-PHOSPHATE	<chem>C(C(C(C(C(=O)O)O)OP(=O)(O)O</chem>
beta-ALANINE	<chem>C(CN)C(=O)O</chem>
L-ALANINE	<chem>CC(C(=O)O)N</chem>

Table S2: Collection of the 20 cofactors used for the chemical space expansion, with their molecule name as well as their input format. Most molecules have abstract labels like [NADH] instead of atom labels for ease of use. CoASH: coenzyme A unbound form, Pi: phosphate, PPi: pyrophosphate, UQred: ubiquitin reduced, UQox: ubiquitin oxidised, Fdred: ferredoxin reduced, Fdox: ferredoxin oxidised.

Molecule Name	Smiles Representation
CoASH	[CoA][S][H]
ATP	[Ad]OP(=O)(O)OP(=O)(O)OP(=O)(O)O
ADP	[Ad]OP(=O)(O)OP(=O)(O)O
AMP	[Ad]OP(=O)(O)O
Pi	OP(=O)(O)O
PPi	OP(=O)(O)OP(=O)(O)O
NADPH	[NADP][H]
NADP+	[NADP+]
NADH	[NAD][H]
NAD+	[NAD+]
Hplus	[H+]
H2O	O
HCO3-	C(=O)(O)[O-]
CO2	C(=O)=O
UQred	O=[UQ]=O
UQox	O[UQ]O
Fdred	[Fdred]
Fdox	[Fdox]
O2	O=O
H2O2	OO

Figure S1: Collection of all graph transformation rules used in the creation of the chemical space on the following 16 pages, including an overview list. The rules are depicted as double pushout diagrams. (L) shows the left side of the reaction, (R) shows the right side of the reaction, and (K) represents the reaction core, with the bonds and labels that are not changing during the reaction. Where applicable, labels like ($_X$), ($_Y$), and ($_Z$), are listed underneath the diagrams with the atom types they can represent. Some labels represent whole cofactors instead of single atoms: NADP(+), NAD(+), UQ: Ubiquitin, Fdox: Ferredoxin oxidised, Fdred: Ferredoxin reduced, CoA: Coenzyme A, Ad: Adenine. The names of the rules are the names of the original enzymes that the rules were modeled after and modified.

Contents

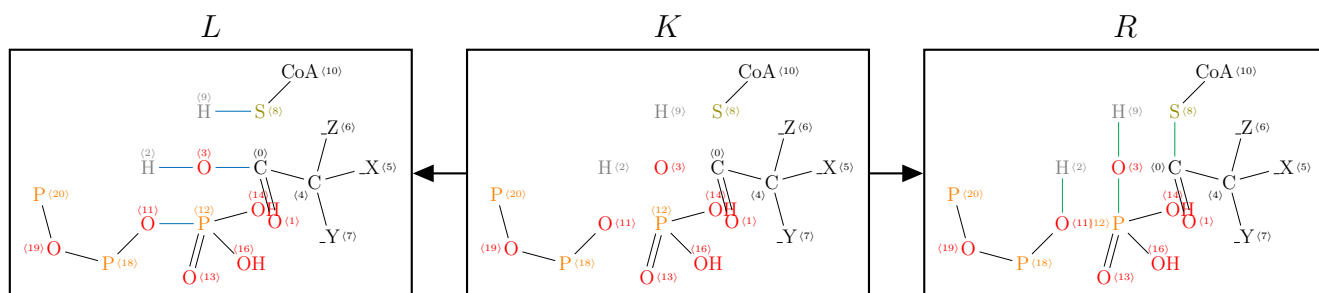
0.1	DPO Diagramms of Rule(s)	3
0.1.1	rTCA_activation_succinyl_CoA_synthetase_01	3
0.1.2	rTCA_activation_ATP_citrate_lyase_01	3
0.1.3	rTCA_reduction_malate_dehydrogenase_01	3
0.1.4	rTCA_reduction_fumarate_reductase_01_2	4
0.1.5	rTCA_dehydration_fumarase_01	4
0.1.6	rTCA_dehydration_rehydration_aconitase_02	4
0.1.7	rTCA_dehydration_rehydration_crotonase_02_1	5
0.1.8	rTCA_carboxylation_2KFOR_PFOR_01	5
0.1.9	rTCA_carboxylation_IDH_01	5
0.1.10	rTCA_carboxylation_PEPC_01_1	6
0.1.11	rTCA_activation_dikinase_01_1	6
0.1.12	3HP4HB_dehydration_4HBCoA_dehydratase_01	6
0.1.13	3HP4HB_carboxylation_acetyl-CoA_carboxylase_01	7
0.1.14	3HP4HB_rearrangement_methylmalonyl-CoA_mutase_01	7
0.1.15	3HP4HB_reduction_acrylyl-CoA_reductase_01	7
0.1.16	3HP4HB_oxidation_3-hydroxybutyryl-CoA_dehydrogenase_01	8
0.1.17	3HP4HB_dehydration_3HPCoA_dehydratase_01	8
0.1.18	3HP4HB_reduction_succinyl-CoA_reductase_01	8
0.1.19	3HP4HB_reduction_3-hydroxypropionate_dehydrogenase_01	9
0.1.20	3HP4HB_activation_4HBCoA_synthetase_01	9
0.1.21	3HP4HB_transferase_acetyl-CoA_C-acyltransferase_01_1	9
0.1.22	3HPbicy_mesaconyl_C1_C4_CoA_transferase_01	10
0.1.23	3HPbicy_activation_methylmalyl-CoA_lyase_01	10
0.1.24	3HPbicy_transferase_succinyl-CoA_malate-CoA_transferase_01	10
0.1.25	3HPbicy_oxidation_succinate_dehydrogenase_01	11
0.1.26	3HPbicy_oxidation_succinate_dehydrogenase_01_1	11
0.1.27	3HPbicy_oxidation_succinate_dehydrogenase_01_2	11
0.1.28	3HPbicy_transferase_succinyl-CoA_malate-CoA_transferase_a_01	12
0.1.29	3HPbicy_transferase_succinyl-CoA_malate-CoA_transferase_b_01	12
0.1.30	CETCH_activation_methylmalyl-CoA_lyase_01	12
0.1.31	CETCH_activation_lysis_malate_synthase_02	13
0.1.32	CETCH_carboxylation_ccr_carboxylase/reductase_01	13
0.1.33	CETCH_dehydration_rehydration_mesCoA_hydratase_02_2	13
0.1.34	CETCH_oxidation_propionyl-CoA/methylsuccinyl-CoA_oxidase_01	14
0.1.35	CETCH_rearrangement_ethylmalonyl-CoA_mutase_02	14
0.1.36	GLYOX_lysis_isocitrate_lyase_01	14
0.1.37	GLYOX_activation_malate_synthase_a_01	15
0.1.38	GLYOX_lysis_isocitrate_lyase_02_reverse	15
0.1.39	GLYOX_alanine_2,3_aminomutase_01	15
0.1.40	GLYOX_transaminase_01	16

0.1.41 GLYOX_transaminase_01 16

0.1.42 GLYOX_methylmalonyl-CoA_carboxytransferase_01 16

0.1 DPO Diagramms of Rule(s)

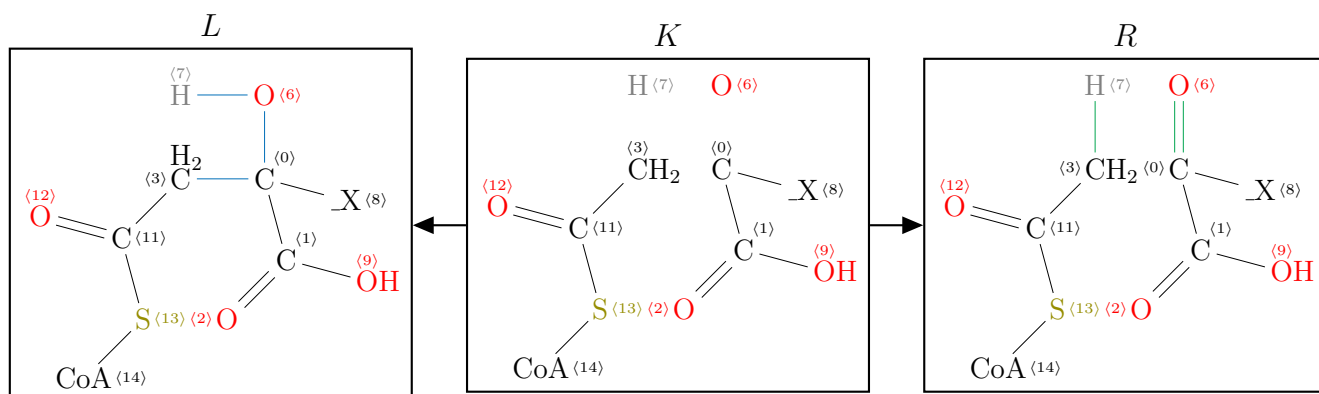
0.1.1 rTCA_activation_succinyl_CoA_synthetase_01



Files: out/001_r_0_11300110_{L, K, R}

$_X \in \{\text{'C'}, \text{'H'}\}$
 $_Y \in \{\text{'C'}, \text{'H'}\}$
 $_Z \in \{\text{'C'}, \text{'H'}\}$

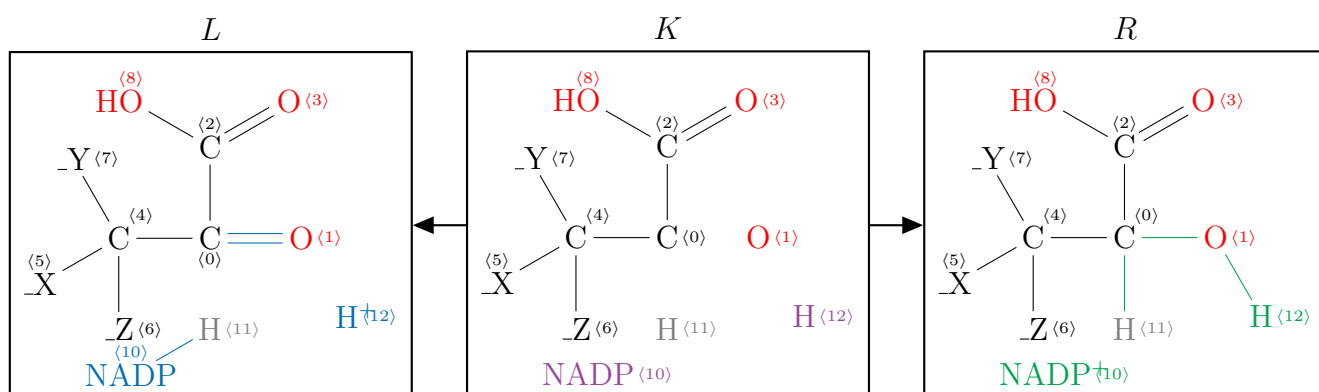
0.1.2 rTCA_activation_ATP_citrate_lyase_01



Files: out/004_r_1_11300110_{L, K, R}

$_X \in \{\text{'C'}, \text{'H'}\}$

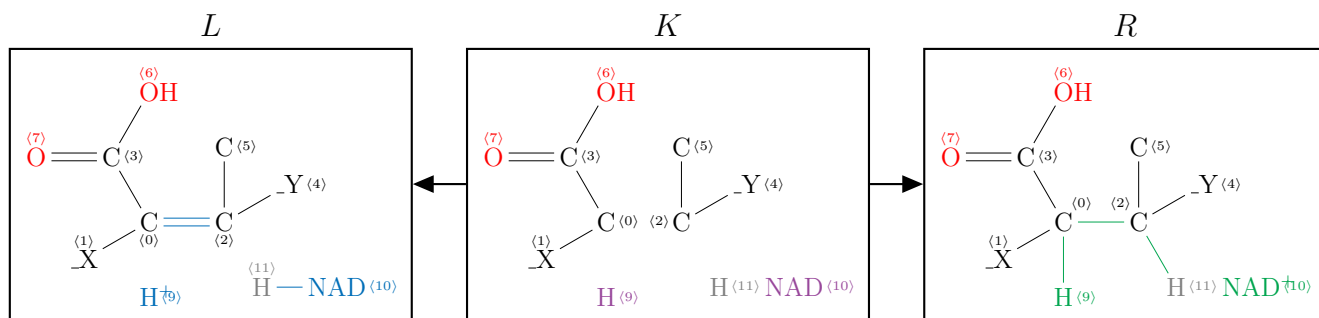
0.1.3 rTCA_reduction_malate_dehydrogenase_01



Files: out/007_r_2_11300110_{L, K, R}

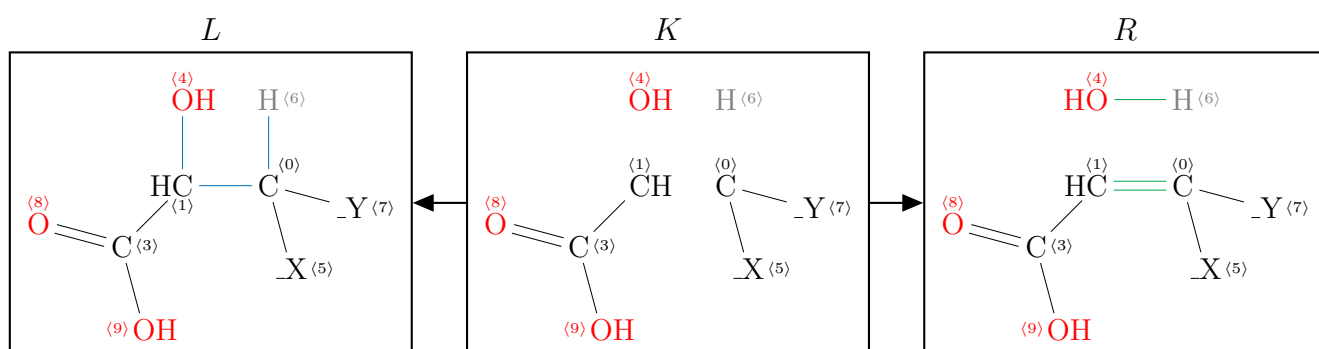
$_X \in \{\text{'C'}, \text{'H'}\}$
 $_Y \in \{\text{'C'}, \text{'H'}\}$
 $_Z \in \{\text{'C'}, \text{'H'}\}$

0.1.4 rTCA_reduction_fumarate_reductase_01_2



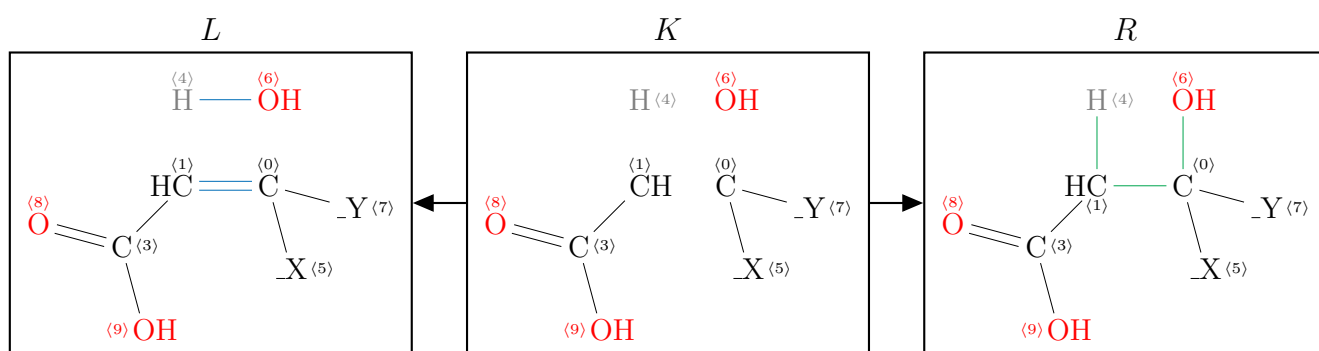
$_X \in \{\text{'H'}, \text{'C'}\}$
 $_Y \in \{\text{'H'}, \text{'C'}\}$

0.1.5 rTCA_dehydration_fumarase_01



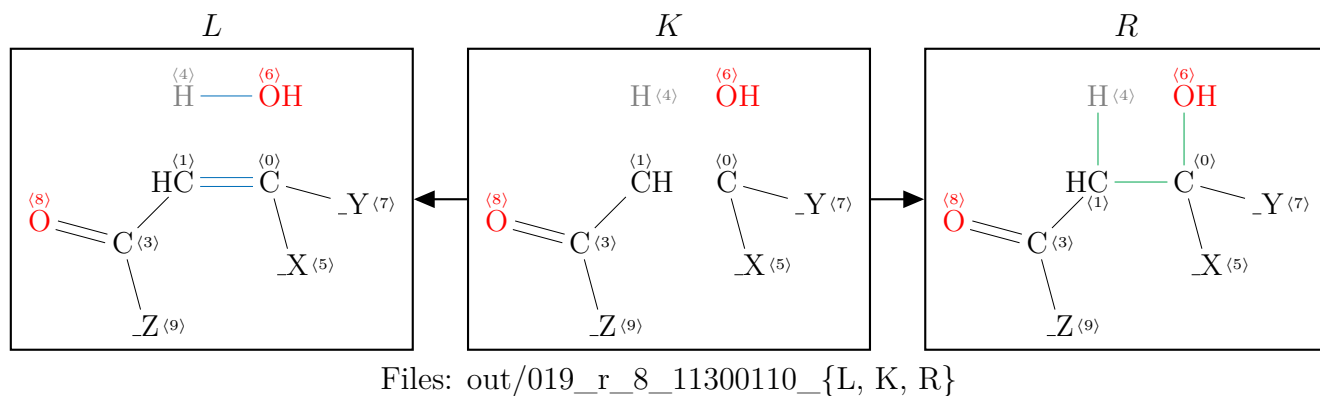
$_X \in \{\text{'C'}, \text{'H'}\}$
 $_Y \in \{\text{'C'}, \text{'H'}\}$

0.1.6 rTCA_dehydration_rehydration_aconitase_02



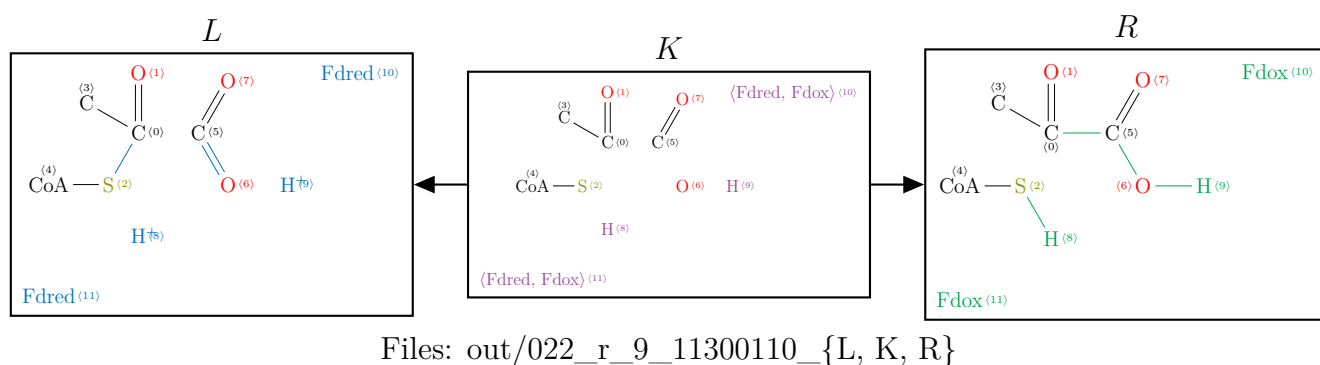
$_X \in \{\text{'C'}\}$
 $_Y \in \{\text{'C'}\}$

0.1.7 rTCA_dehydration_rehydration_crotonase_02_1

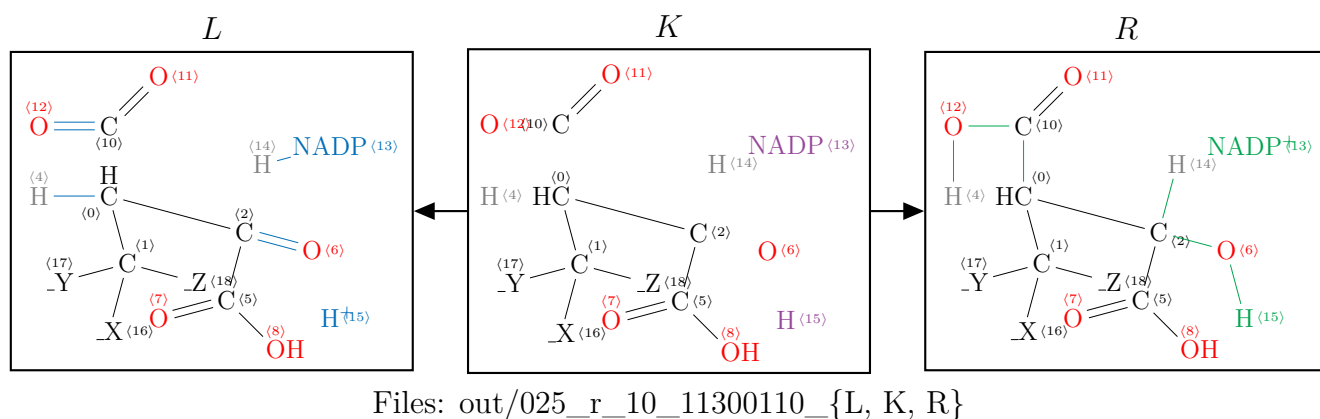


$_X \in \{\text{'C'}, \text{'H'}\}$
 $_Y \in \{\text{'C'}, \text{'H'}\}$
 $_Z \in \{\text{'O'}, \text{'S'}\}$

0.1.8 rTCA_carboxylation_2KFOR_PFOR_01

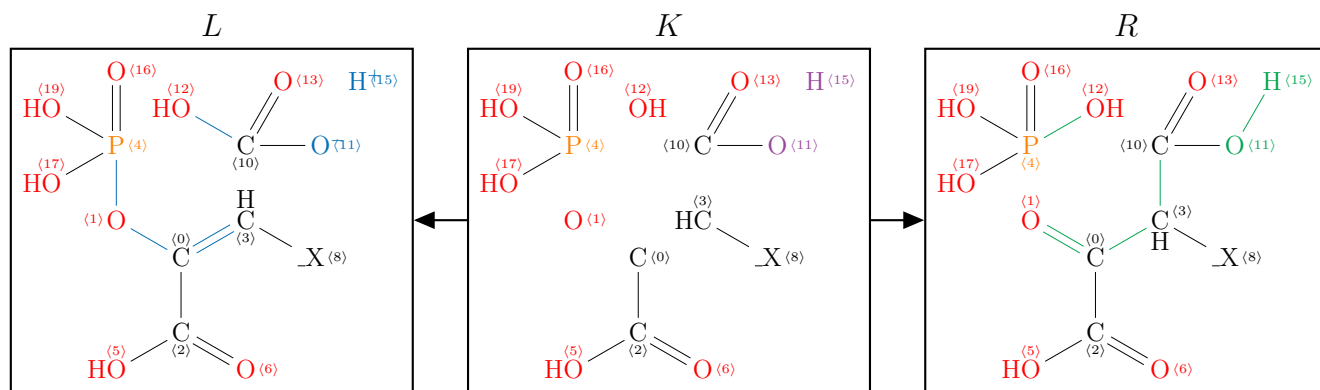


0.1.9 rTCA_carboxylation_IDH_01



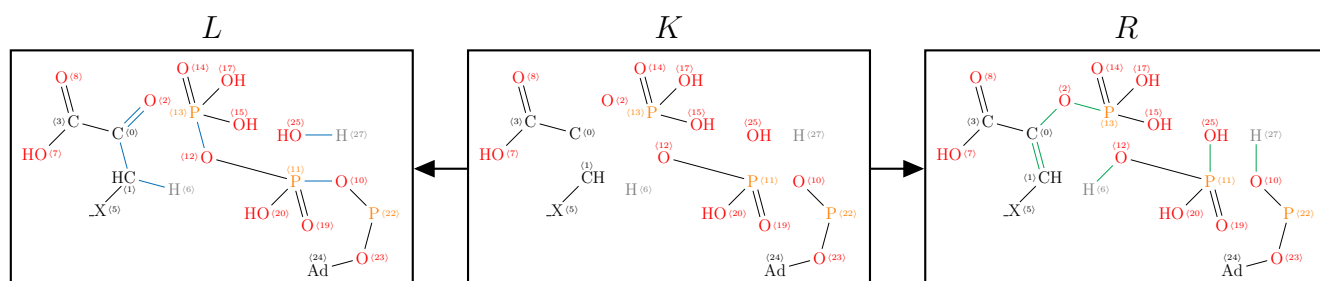
$_X \in \{\text{'C'}, \text{'H'}\}$
 $_Y \in \{\text{'C'}, \text{'H'}\}$
 $_Z \in \{\text{'C'}, \text{'H'}\}$

0.1.10 rTCA_carboxylation_PEPC_01_1



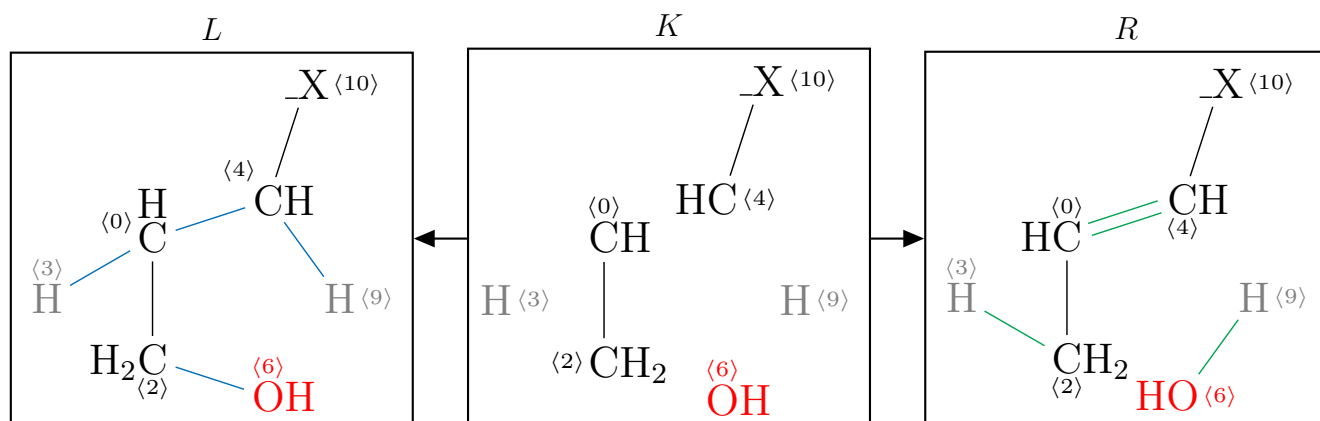
$_X \in \{\text{'H'}, \text{'C'}\}$

0.1.11 rTCA_activation_dikinase_01_1



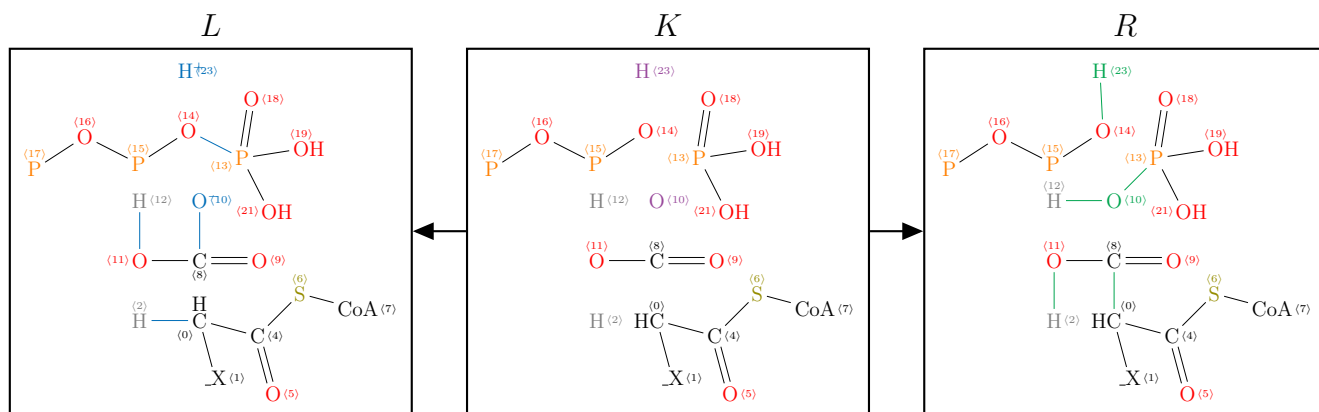
$_X \in \{\text{'H'}, \text{'C'}\}$

0.1.12 3HP4HB_dehydration_4HBCoA_dehydratase_01



$_X \in \{\text{'C'}\}$

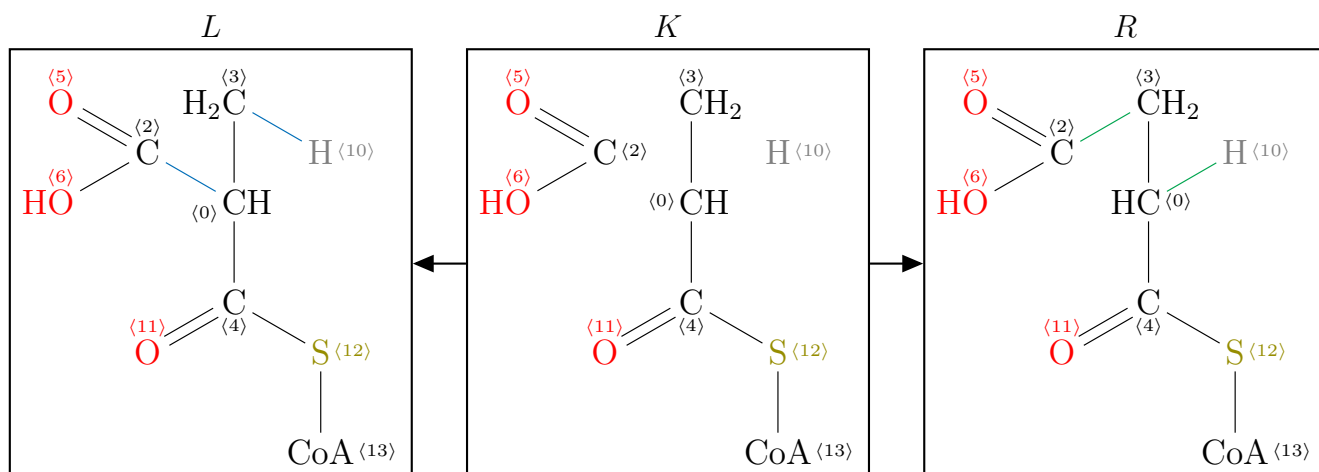
0.1.13 3HP4HB_carboxylation_acetyl-CoA_carboxylase_01



Files: out/037_r_16_11300110_{L, K, R}

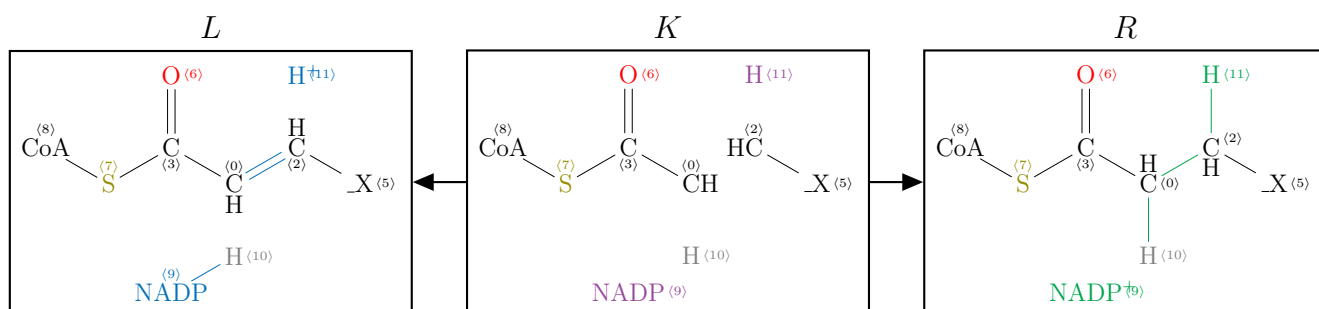
$_X \in \{\text{'C'}, \text{'H'}\}$

0.1.14 3HP4HB_rearrangement_methylmalonyl-CoA_mutase_01



Files: out/040_r_17_11300110_{L, K, R}

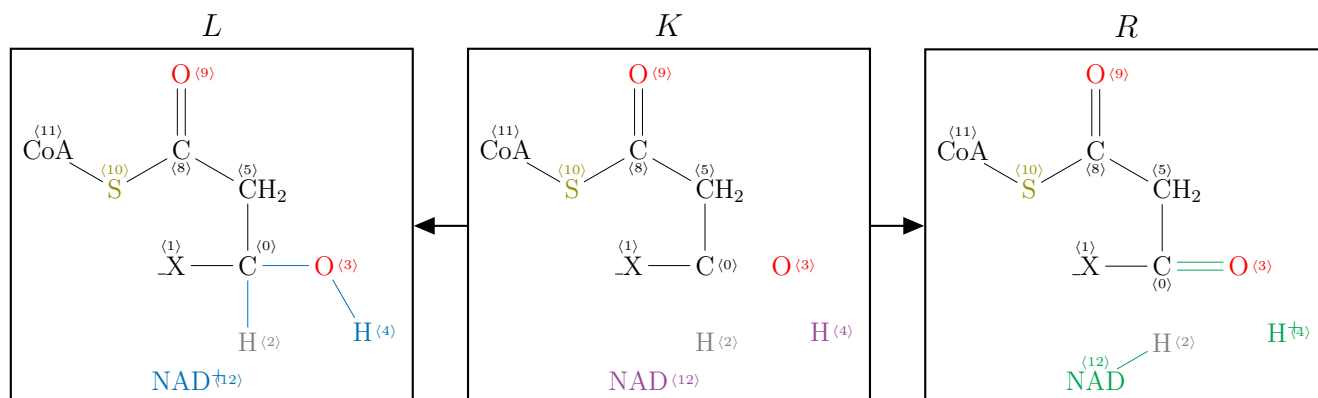
0.1.15 3HP4HB_reduction_acrylyl-CoA_reductase_01



Files: out/043_r_18_11300110_{L, K, R}

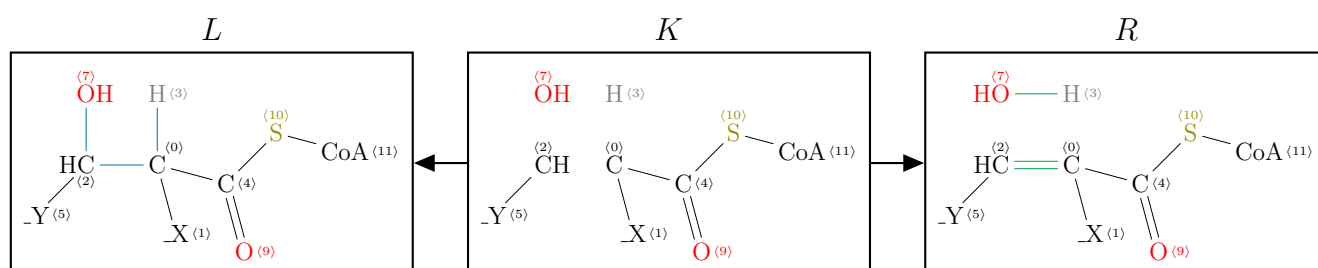
$_X \in \{\text{'C'}, \text{'H'}\}$

0.1.16 3HP4HB_oxidation_3-hydroxybutyryl-CoA_dehydrogenase_01



$_X \in \{\text{'C'}, \text{'H'}\}$

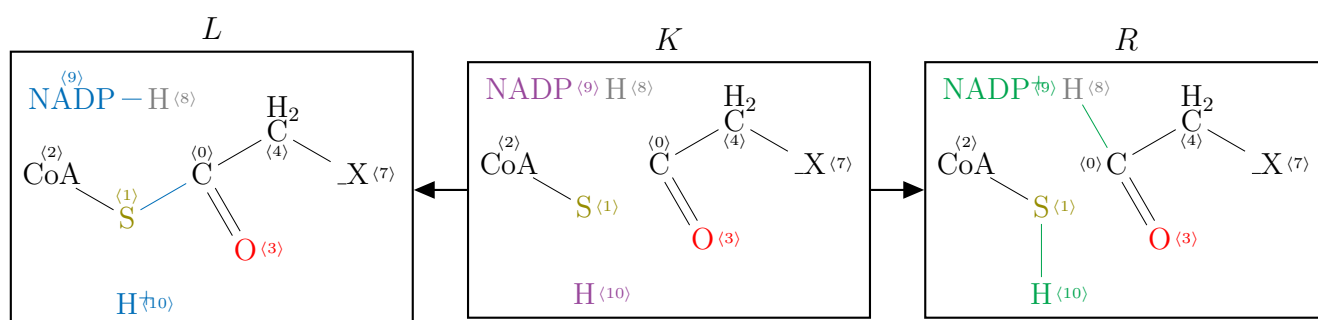
0.1.17 3HP4HB_dehydration_3HPCoA_dehydratase_01



$_X \in \{\text{'C'}, \text{'H'}\}$

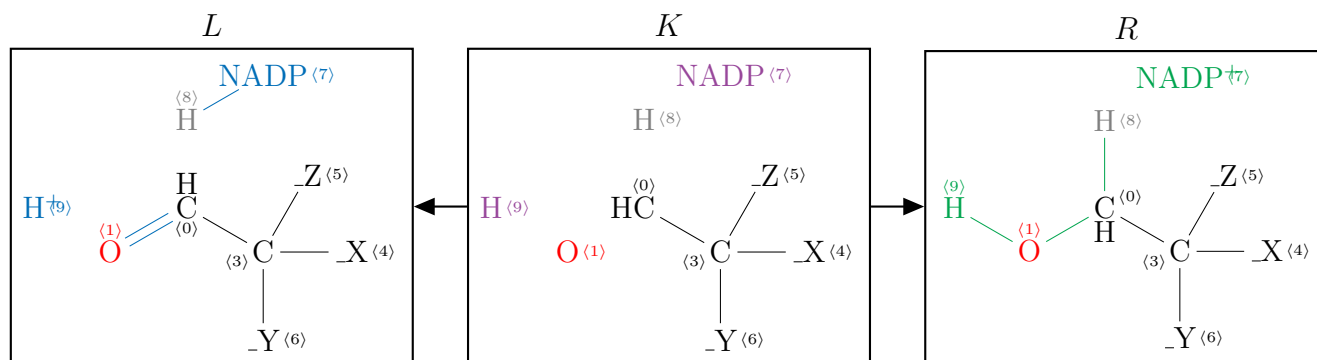
$_Y \in \{\text{'C'}, \text{'H'}\}$

0.1.18 3HP4HB_reduction_succinyl-CoA_reductase_01



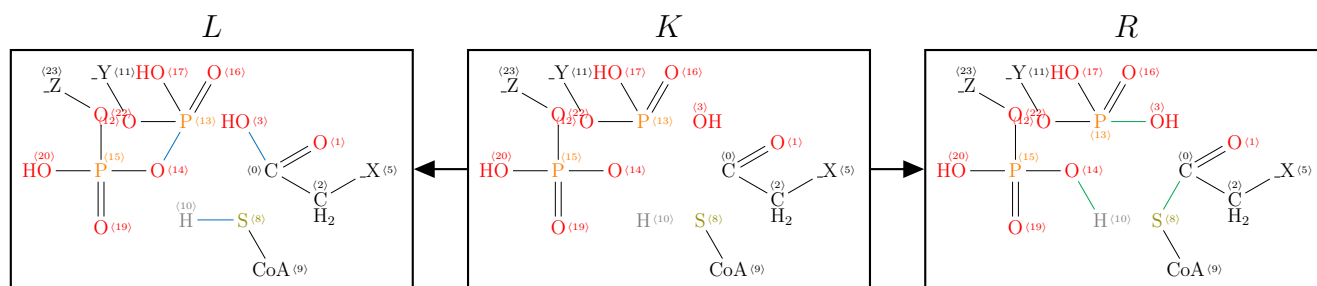
$_X \in \{\text{'C'}\}$

0.1.19 3HP4HB_reduction_3-hydroxypropionate_dehydrogenase_01



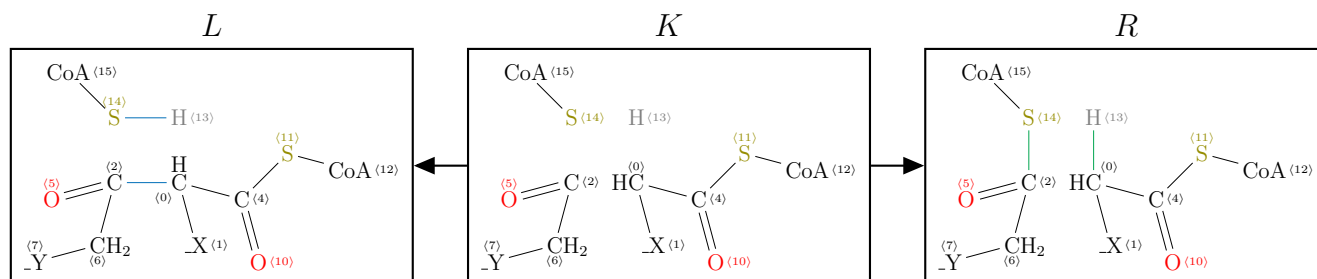
$_X \in \{\text{'C'}, \text{'H'}\}$
 $_Y \in \{\text{'C'}, \text{'H'}\}$
 $_Z \in \{\text{'C'}, \text{'H'}\}$

0.1.20 3HP4HB_activation_4HBCoA_synthetase_01



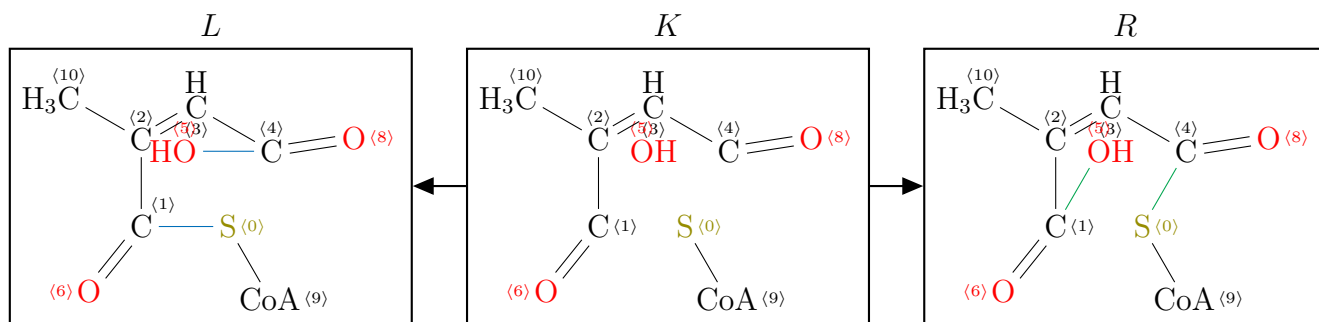
$_X \in \{\text{'C'}\}$
 $_Y \in \{\text{'H'}, \text{'P'}\}$
 $_Z \in \{\text{'P'}, \text{'Ad'}\}$

0.1.21 3HP4HB_transferase_acetyl-CoA_C-acyltransferase_01_1



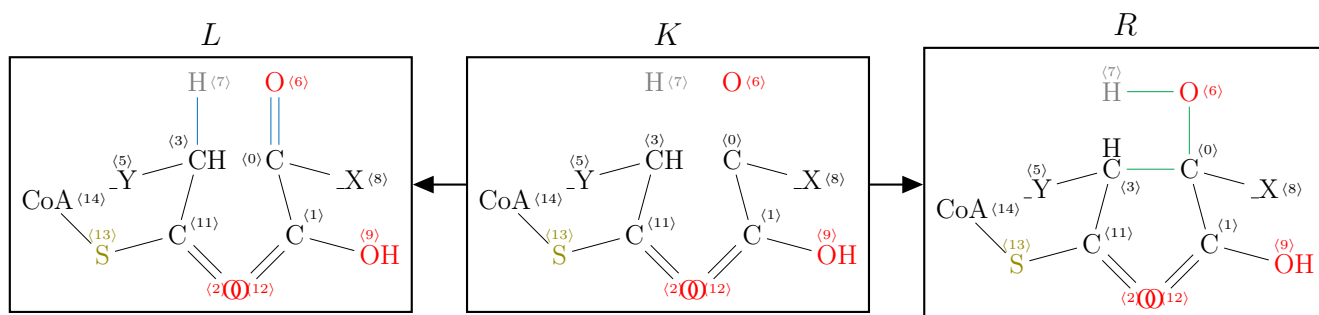
$_X \in \{\text{'H'}, \text{'C'}\}$
 $_Y \in \{\text{'H'}, \text{'C'}\}$

0.1.22 3HPbicy_mesaconyl_C1_C4_CoA_transferase_01



Files: out/064_r_26_11300110_{L, K, R}

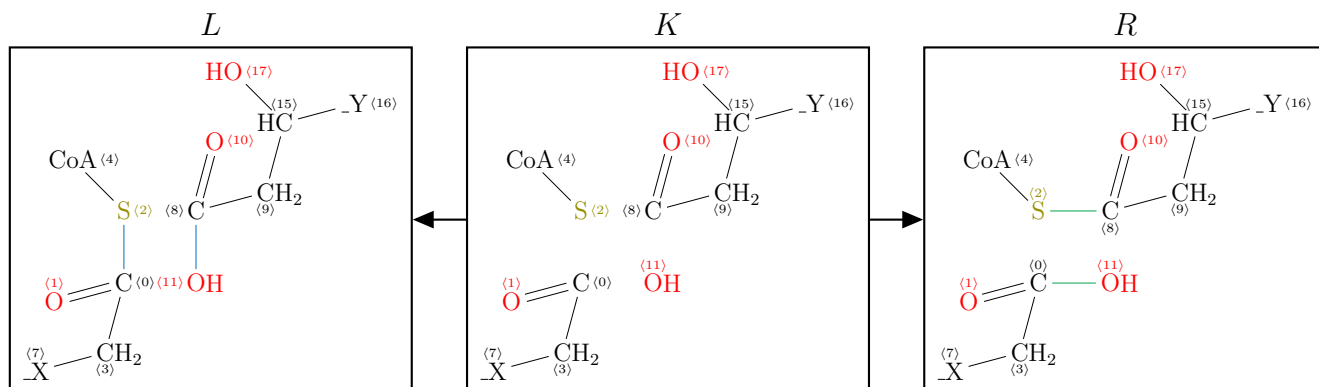
0.1.23 3HPbicy_activation_methylmalyl-CoA_lyase_01



Files: out/067_r_27_11300110_{L, K, R}

$_X \in \{'C', 'H'\}$
 $_Y \in \{'C', 'H'\}$

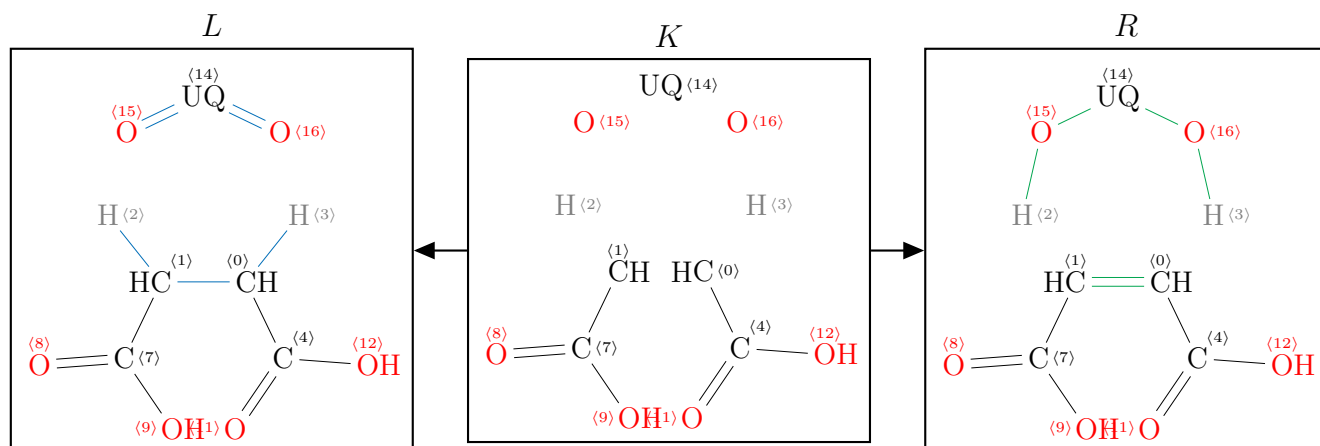
0.1.24 3HPbicy_transferase_succinyl-CoA_malate-CoA_transferase_01



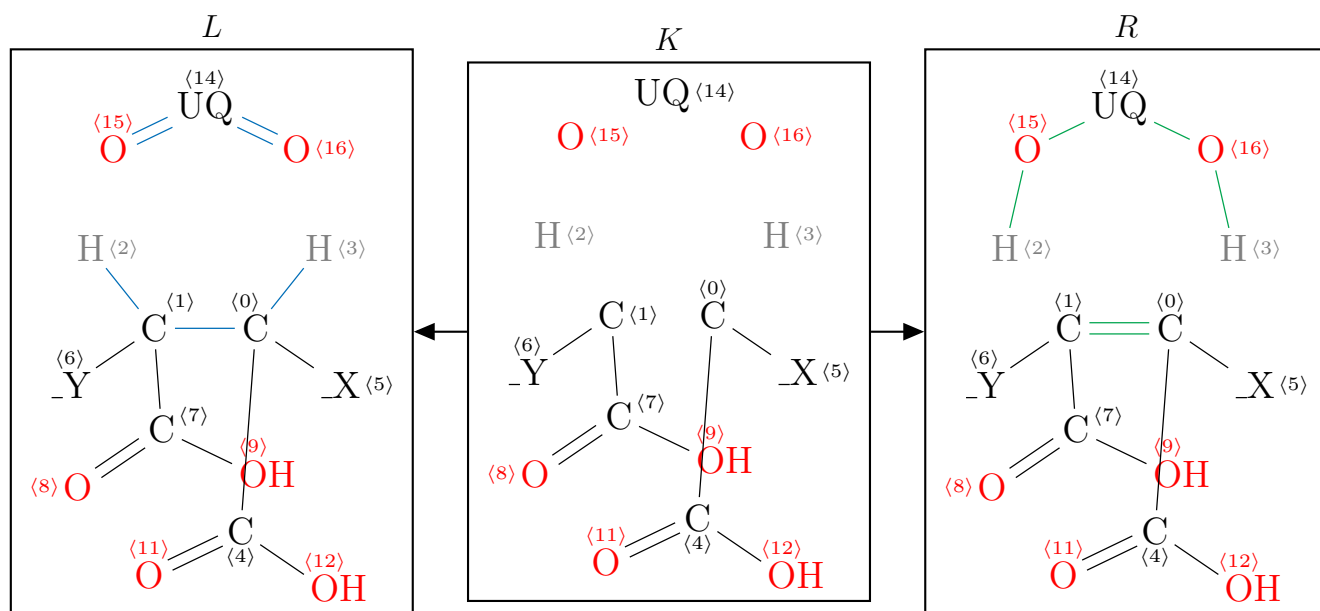
Files: out/070_r_28_11300110_{L, K, R}

$_X \in \{'C'\}$
 $_Y \in \{'C'\}$

0.1.25 3HPbicy_oxidation_succinate_dehydrogenase_01

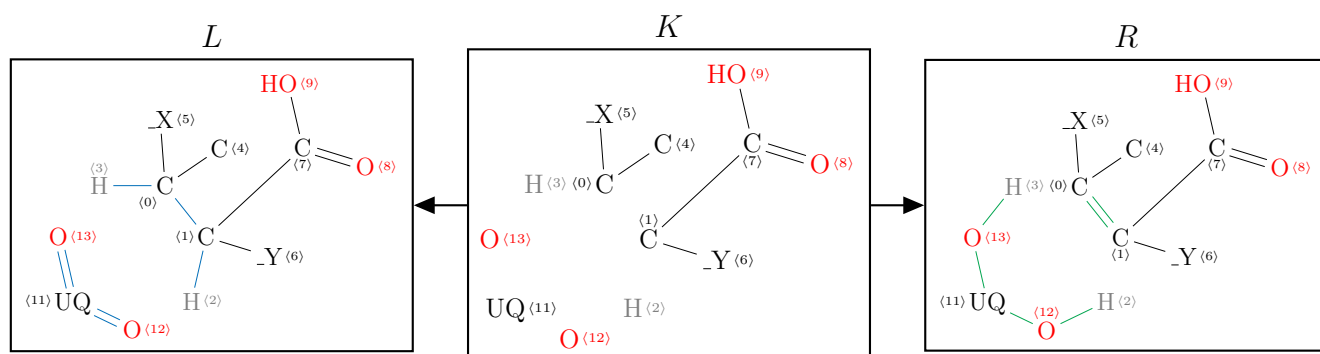


0.1.26 3HPbicy_oxidation_succinate_dehydrogenase_01_1



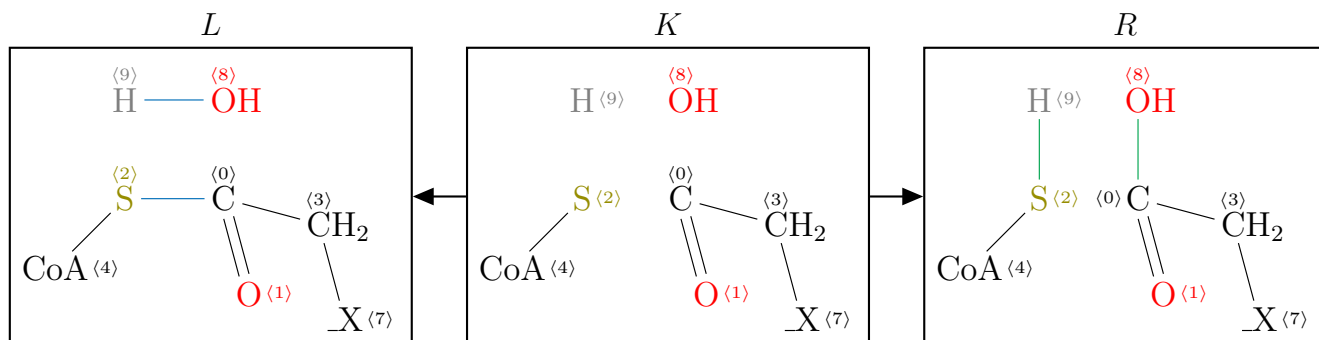
$_X \in \{\text{'H'}, \text{'C'}\}$
 $_Y \in \{\text{'H'}, \text{'C'}\}$

0.1.27 3HPbicy_oxidation_succinate_dehydrogenase_01_2



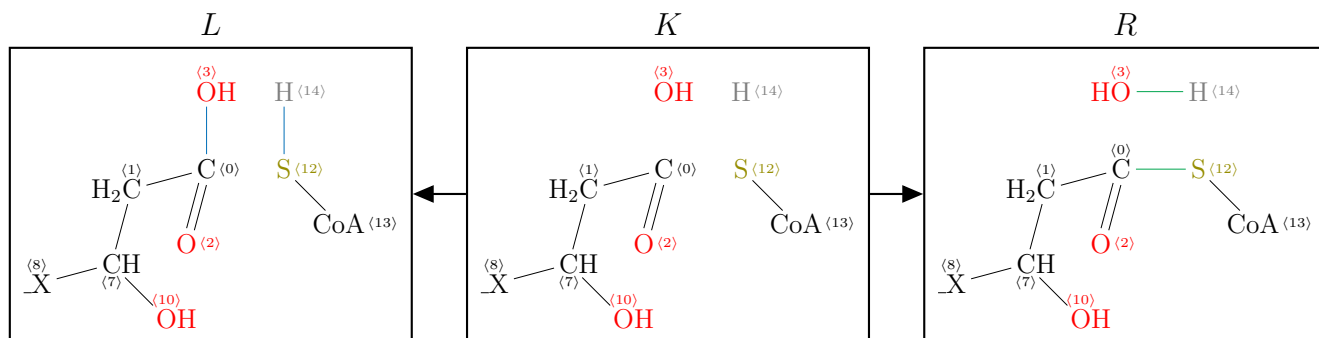
$_X \in \{\text{'H'}, \text{'C'}\}$
 $_Y \in \{\text{'H'}, \text{'C'}\}$

0.1.28 3HPbicy_transferase_succinyl-CoA_malate-CoA_transferase_a_01



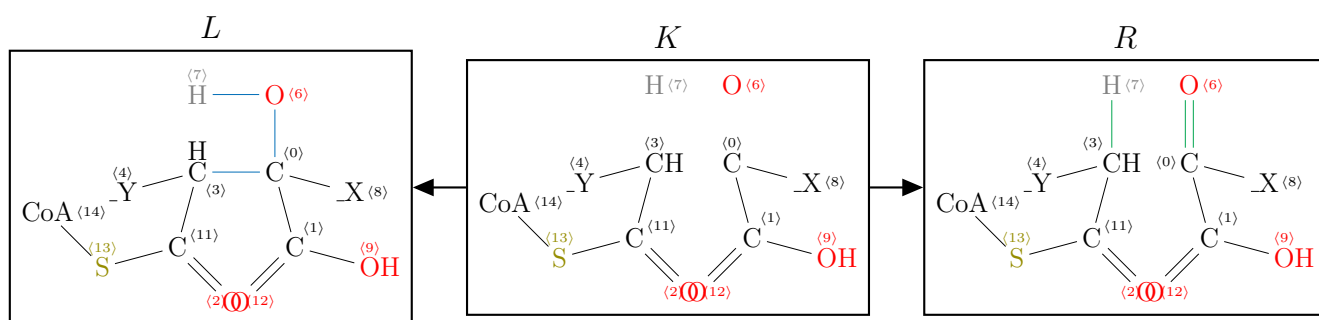
$_X \in \{\text{'H'}, \text{'C'}\}$

0.1.29 3HPbicy_transferase_succinyl-CoA_malate-CoA_transferase_b_01



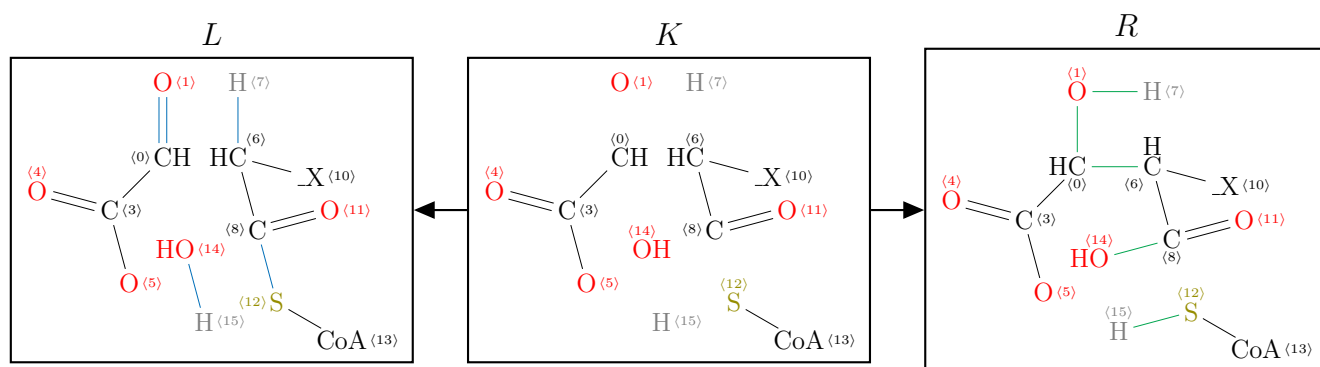
$_X \in \{\text{'H'}, \text{'C'}\}$

0.1.30 CETCH_activation_methylmalyl-CoA_lyase_01



$_Y \in \{\text{'C'}, \text{'H'}\}$
 $_X \in \{\text{'C'}, \text{'H'}\}$

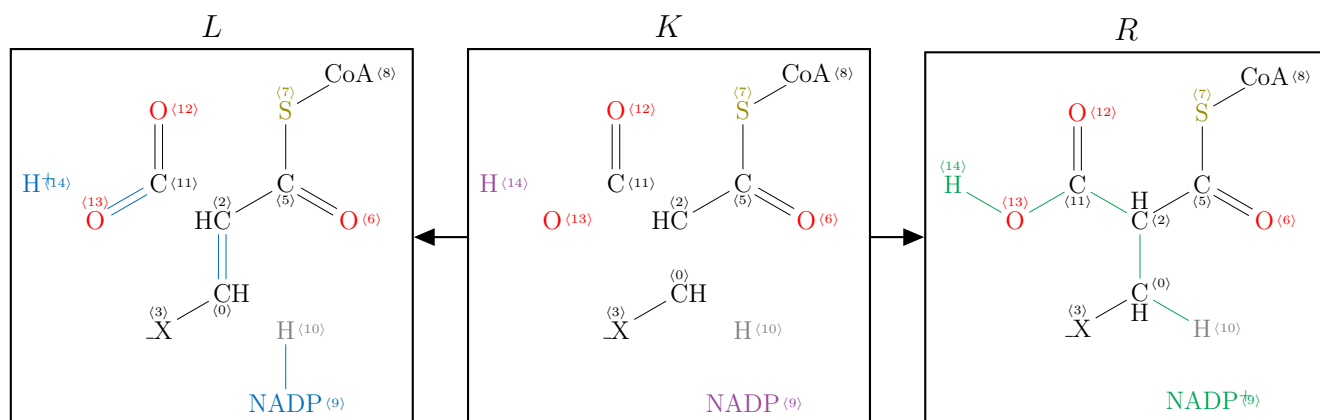
0.1.31 CETCH_activation_lysis_malate_synthase_02



Files: out/091_r_47_11300110_{L, K, R}

$_X \in \{\text{'C'}, \text{'H'}\}$

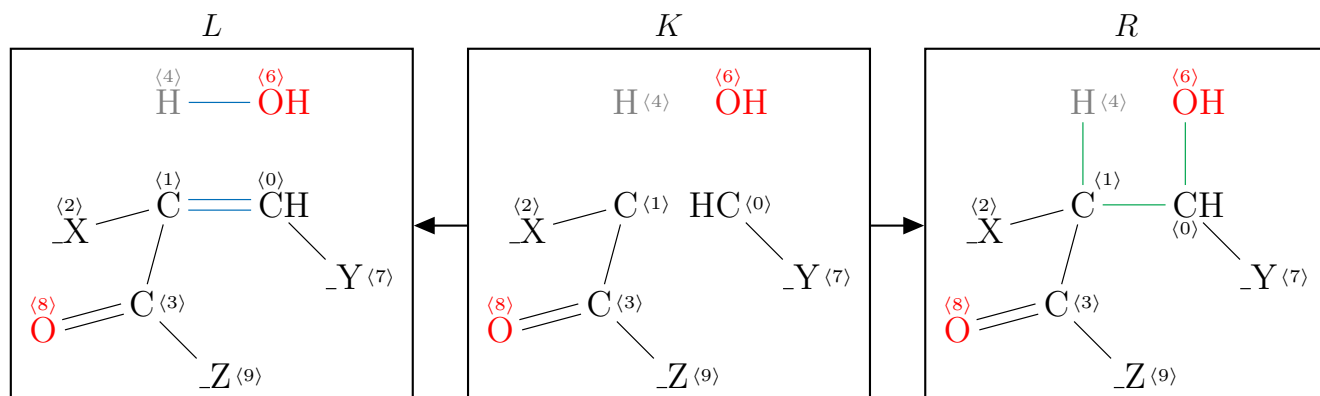
0.1.32 CETCH_carboxylation_ccr_carboxylase/reductase_01



Files: out/094_r_48_11300110_{L, K, R}

$_X \in \{\text{'C'}, \text{'H'}\}$

0.1.33 CETCH_dehydration_rehydration_mesCoA_hydratase_02_2



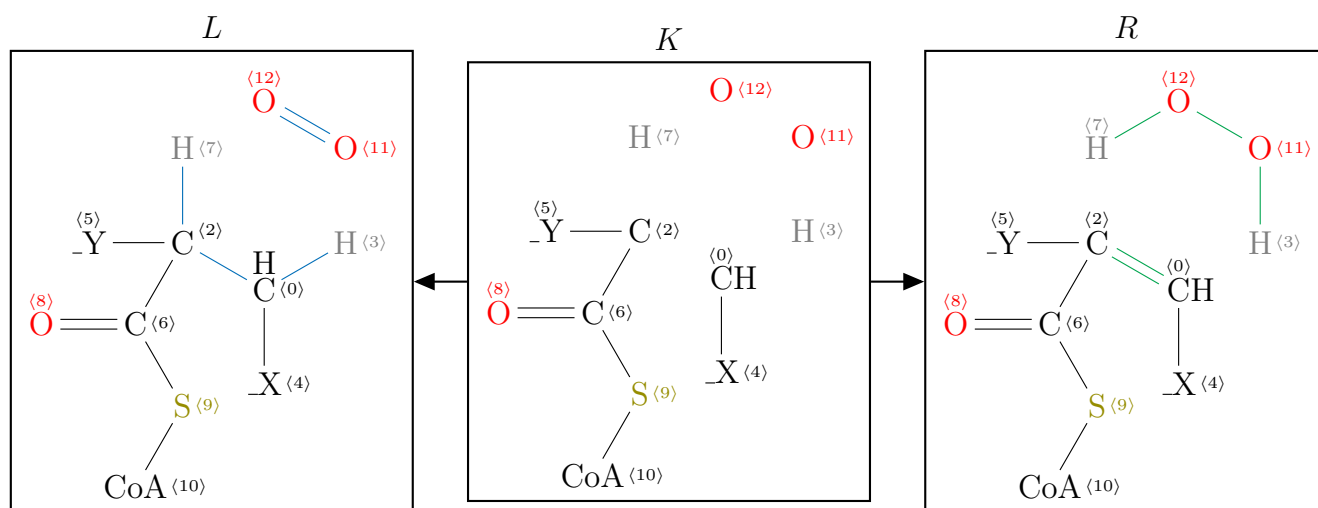
Files: out/097_r_49_11300110_{L, K, R}

$_X \in \{\text{'C'}, \text{'H'}\}$

$_Y \in \{\text{'C'}, \text{'H'}\}$

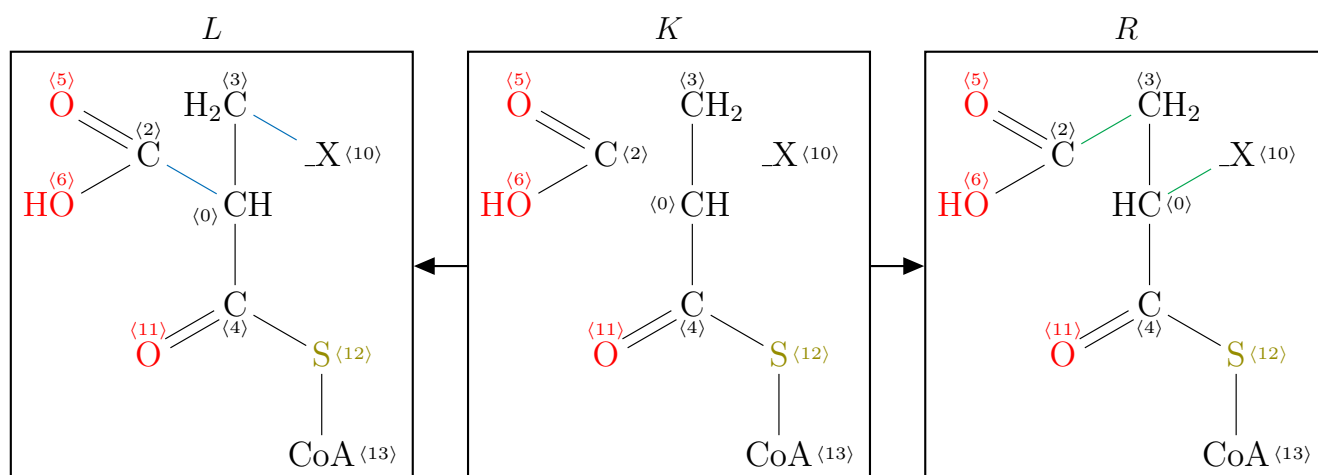
$_Z \in \{\text{'O'}, \text{'S'}\}$

0.1.34 CETCH_oxidation_propionyl-CoA/methylsuccinyl-CoA_oxidase_01



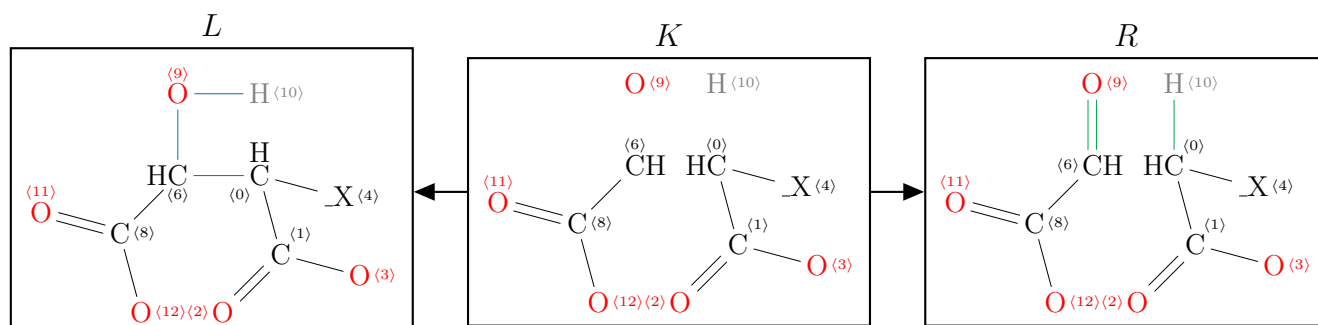
$_X \in \{\text{'C'}, \text{'H'}\}$
 $_Y \in \{\text{'C'}, \text{'H'}\}$

0.1.35 CETCH_rearrangement_ethylmalonyl-CoA_mutase_02



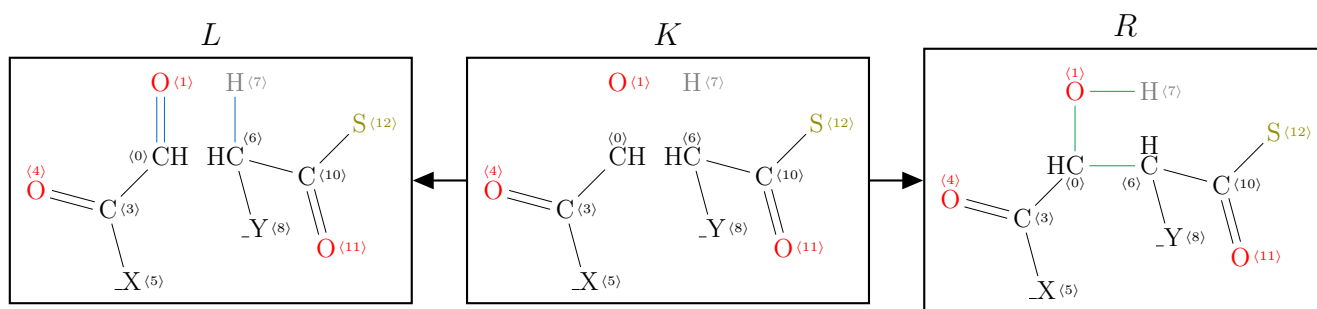
$_X \in \{\text{'C'}, \text{'H'}\}$

0.1.36 GLYOX_lysis_isocitrate_lyase_01



$_X \in \{\text{'H'}, \text{'C'}\}$

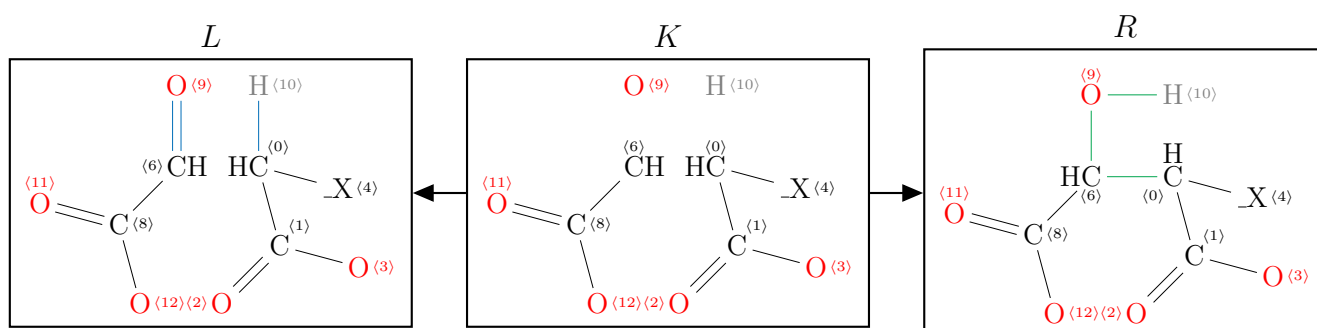
0.1.37 GLYOX_activation_malate_synthase_a_01



Files: out/109_r_53_11300110_{L, K, R}

$_X \in \{'O', 'S'\}$
 $_Y \in \{'C', 'H'\}$

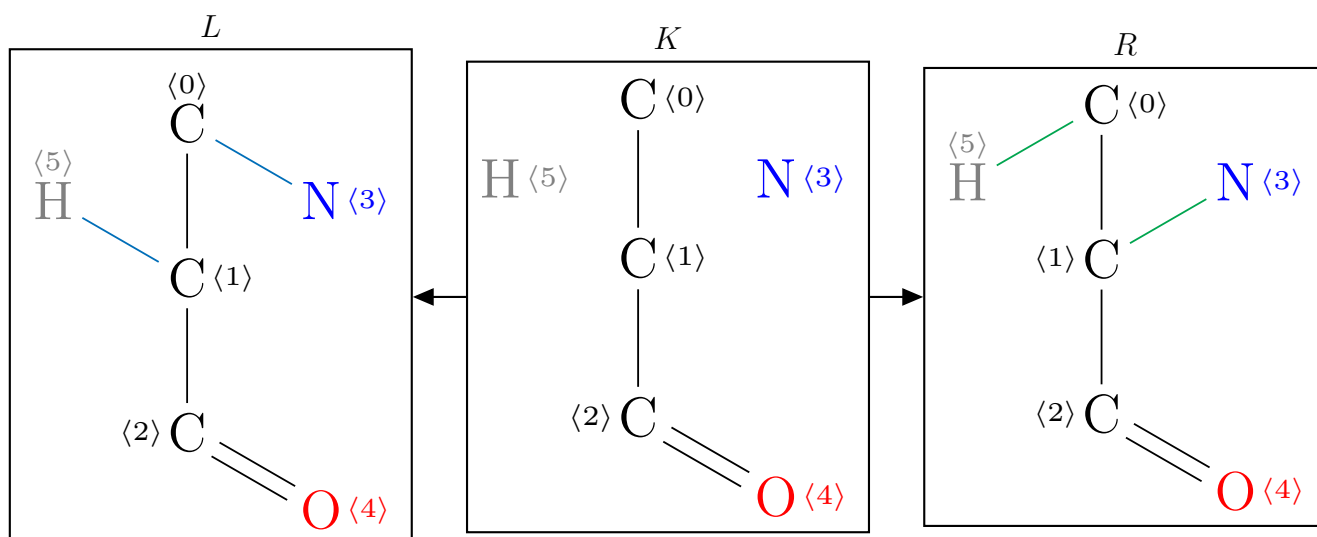
0.1.38 GLYOX_lysis_isocitrate_lyase_02_reverse



Files: out/112_r_54_11300110_{L, K, R}

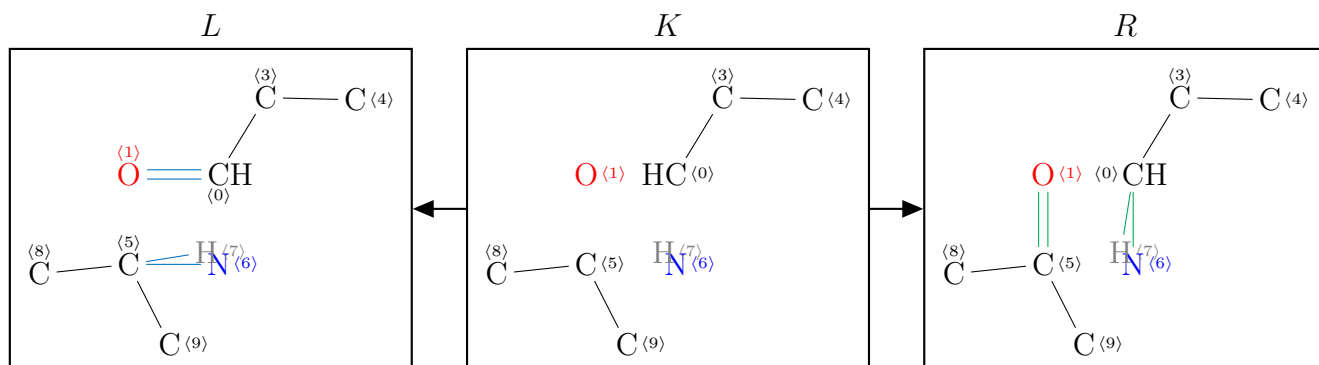
$_X \in \{'H', 'C'\}$

0.1.39 GLYOX_alanine_2,3_aminomutase_01

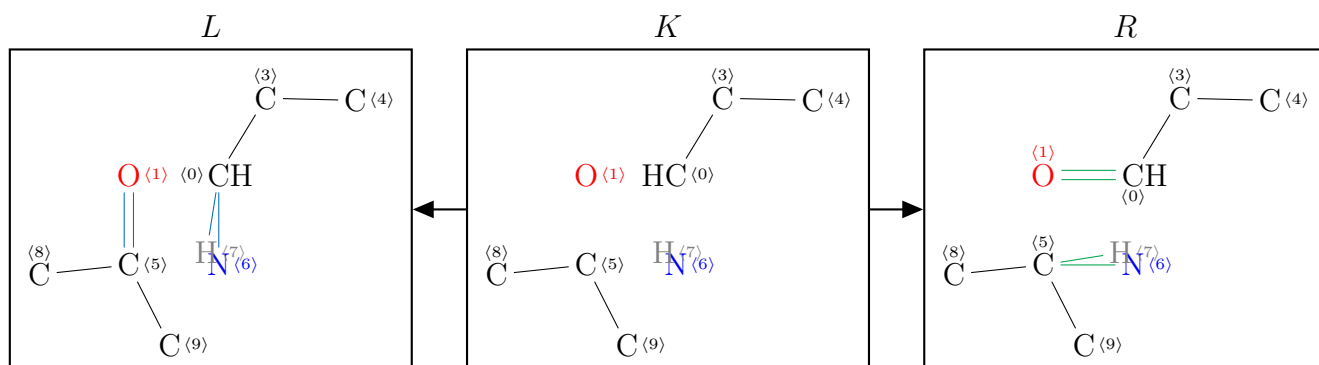


Files: out/115_r_55_11300110_{L, K, R}

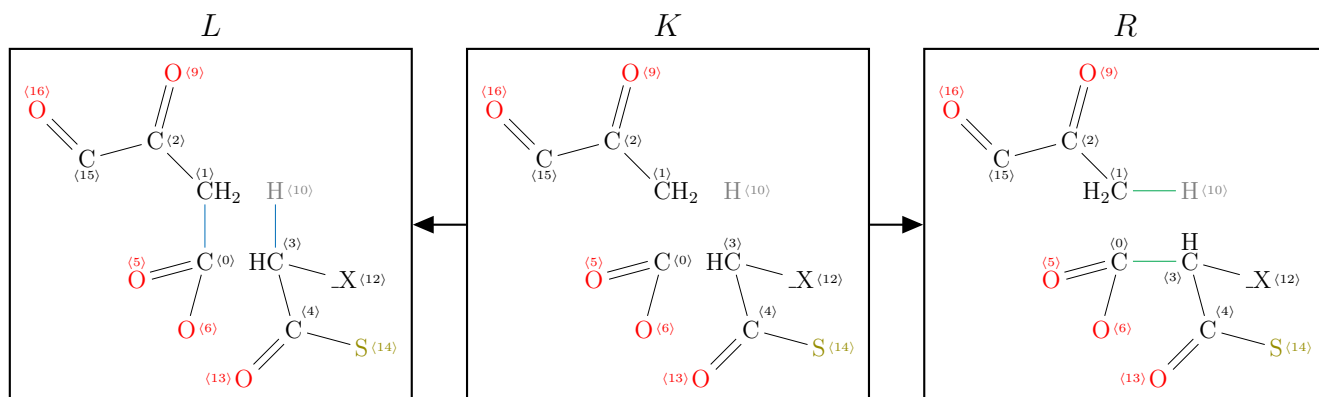
0.1.40 GLYOX_transaminase_01



0.1.41 GLYOX_transaminase_01



0.1.42 GLYOX_methylmalonyl-CoA_carboxytransferase_01



$_X \in \{\text{'H'}, \text{'C'}\}$